Conceitos básicos de estatística

META

Apresentar conceitos básicos de estatística necessários para o estudo da mecânica estatística.

OBJETIVOS

Ao fim da aula os alunos deverão ser capazes de:

Entender conceitos básicos de estatística como espaço amostral, eventos e probabilidade.

Compreender as regras da soma e da multiplicação da probabilidade.

Entender os conceitos de variáveis aleatórias, densidade de probabilidade e seus momentos.

Resolver problemas envolvendo estes conceitos.

PRÉ-REQUISITOS

Cálculo diferencial e integral, teoria de conjuntos elementar e função δ de Dirac.

4.1 Introdução

Prezado aluno, antes de explorarmos a mecânica estatística, temos que entender alguns conceitos básicos de estatística. Esta é uma ciência que estuda a coleção, organização e a interpretação de dados. A estatística pode ser basicamente dividida em dois ramos:

Estatística indutiva: produzir afirmações sobre uma dada característica da população, na qual estamos interessados, a partir de informações colhidas de uma parte dessa população.

Estatística descritiva: reúne um conjunto de técnicas para sumarizar os dados e medidas descritivas que permitem tirar muitas informações contidas nos dados.

No estudo de mecânica estatística estaremos interessados na estatística descritiva, pois é através desta que buscaremos técnicas para extrair a descrição macroscópica de um sistema por meio da observação e sumarização de dados microscópicos dos seus constituintes.

4.2 Probabilidade

O conceito de probabilidade está ligado à quantificação da possibilidade de ocorrência de um certo evento. Iremos usar a estratégia de representar eventos através de conjuntos*. Isto nos permitirá usar inúmeras regras da teoria de conjuntos que você já conhece desde o ensino médio. Para iniciarmos o estudo de probabilidade, vamos apresentar algumas definições:

*Teoria de conjuntos é apenas uma maneira para tratar eventos em probabilidade. Existem outras maneiras de fazer este tratamento como, por exemplo, representar um evento como uma variável lógica (verdadeiro ou falso) e tratar esta variável com a álgebra de Boole. Maiores detalhes na ref. [2].

Experimento: qualquer processo que nos permita obter observações. Um experimento pode ser classificado como:

- Determinístico: quando o resultado do experimento já é conhecido. Ex.: jogar uma pedra para cima e observar se ela cai ou não.

- Aleatório: quando o resultado do experimento não é conhecido, mesmo sabendo todas as possibilidades. Ex.: jogar uma moeda para o alto e observar a face (cara ou coroa) que cai para cima.

Espaço amostral: é o conjunto de todos resultados possíveis para um experimento aleatório. Ex.: para o lançamento de um dado, o espaço amostral é $\Omega = \{1, 2, 3, 4, 5, 6\}$, onde cada um destes números representa a face do dado que cai voltada para cima.

Evento: é qualquer subconjunto do espaço amostral. Ex.: O conjunto $\{3, 5\}$ é um evento do lançamento de dados pois está contido em Ω. O conjunto $\{1\}$ é um evento simples do lançamento de dados pois, além de estar contido em Ω, não pode ser particionado.

Eventos mutualmente excludentes: São eventos que quando um ocorre, o outro não pode ocorrer. Isto implica que, se A e B são mutualmente excludentes, então $A \cap B = \emptyset$, ou seja, os conjuntos A e B são disjuntos. Ex.: Considere um jogo de dado onde A seja o evento de ocorrência de faces com números maiores que dois, B seja o evento de occorência de faces menores que três e C seja o evento de ocorrência de faces com números primos. Sendo assim, $A = \{3, 4, 5, 6\}$,

$B = \{1, 2\}$ e $C = \{2, 3, 5\}$. Note que

$$\begin{aligned} A \cap B &= \emptyset \Leftrightarrow A \text{ e } B \text{ são mutualmente excludentes.} \\ A \cap C &= \{3, 5\} \Leftrightarrow A \text{ e } C \text{ não são mutualmente excludentes.} \\ B \cap C &= \{2\} \Leftrightarrow B \text{ e } C \text{ não são mutualmente excludentes.} \end{aligned}$$

Eventos independentes: Se o evento A não interfere na ocorrência do evento B e vice-versa, A e B são chamados de eventos independentes. Ex.: considere o primeiro lançamento de um dado e a ocorrência de uma determinada face (primeiro evento). Ao lançar o dado pela segunda vez, o resultado do evento obtido no primeiro lançamento não influencia o resultado de evento do segundo lançamento. Por isso, podemos dizer que a cada lançamento o dado produz eventos independentes.

Evento complementar: considere um evento $A \subset \Omega$. O evento complementar a este evento é $\tilde{A}$, tal que

$$\tilde{A} \subset \Omega, \quad A \cap \tilde{A} = \emptyset \quad A \cup \tilde{A} = \Omega,$$

ou seja, $\tilde{A}$ contem todos os eventos simples que estão em Ω e não estão em A. Este evento também é chamado de evento negação de A. Ex.: Se para um jogo de dado o evento A é referente a ocorrência de faces com números pares, $A = \{2, 4, 6\}$, então $\tilde{A} = \{1, 3, 5\}$, o qual é referente às faces de números ímpares.

Não existe um acordo bem estabelecido para a atual definição de probabilidade [2]. No entanto, três axiomas* devem ser respeitados:

Axioma I: A probabilidade de um evento A é um número real e não-negativo, ou seja,

$$P(A) \geq 0.$$

Axioma II: A probabilidade de um evento que certamente ocorre é igual à unidade, ou seja,

$$P(\Omega) = 1.$$

O evento que certamente ocorre corresponde ao espaço amostral.

Axioma III: A probabilidade de ocorrer qualquer um dos N eventos mutualmente exclusivos, $A_1, A_2, \ldots, A_N$, é igual a soma da probabilidade de ocorrer cada um deles, ou seja,

$$P(A_1 \cup A_2 \cup \ldots \cup A_N) = P(A_1) + P(A_2) + ... + P(A_N)$$

Estes axiomas são denominados de axiomas de Komolgorov.

Observando várias teorias e escolas de pensamento, duas abordagens emergem de maneira mais frequente. Iremos apresentá-las através de duas regras para encontrar probabilidades:

Abordagem frequencista (empírica): Realize um experimento inúmeras vezes e conte quantas vezes um evento A ocorre. A probabilidade deste evento é estimada por

$$P(A) = \frac{\text{número de vezes que } A \text{ ocorreu}}{\text{número de vezes que o experimento foi realizado}},$$

quando o número de realizações do experimento tende ao infinito.

Abordagem clássica: Considere um experimento onde existem l eventos simples distintos, os quais possuem a mesma chance de acontecer. Se A pode ocorrer s destas l maneiras, então

$$P(A) = \frac{s}{l},$$

ou de maneira análoga

$$P(A) = \frac{n(A)}{n(\Omega)}, \tag{4.1}$$

onde n denota o número de elementos do conjunto.

Exemplo 4.2.1. Calcule, através da abordagem clássica, a probabilidade de ocorrer:
(a) a face 3 no lançamento de um dado "honesto";
(b) o valor 4 na soma do número das faces no lançamento de dois dados "honestos".

Solução: A palavra "honesto" se refere a um dado não viciado, ou seja, um dado onde a probabilidade de ocorrer qualquer face é a mesma.

(a) O espaço amostral no lançamento de um único dado é

$$\Omega = \{1, 2, 3, 4, 5, 6\},$$

No entanto, o evento que ocorre a face 3 é um evento simples, o qual pode ser representado da seguinte forma

$$A = \{3\}.$$

Com isso, podemos calcular a probabilidade deste evento acontecer usando a eq. (4.1)

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{1}{6}$$

(b) O espaço amostral no lançamento de dois dados é

$$\Omega = \quad \{(i, j) \mid i, j = 1, 2, 3, 4, 5, 6\} \tag{4.2}$$

onde i é referente à face do primeiro dado e j para o segundo. Note que $n(\Omega) = 36$. Por outro lado, o evento em que a soma do número das faces é igual a 4 é

$$A = \{(1, 3); (2, 2); (3, 1)\}.$$

Portanto, a probabilidade deste evento ocorrer é

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{36} = \frac{1}{12}.$$

4.2.1 Regra da soma

Qual a probabilidade de acontecer um evento A ou um evento B, tal que $A \subset \Omega$ e $B \subset \Omega$? Para responder esta questão podemos fazer uso da teoria de conjuntos com o número de elementos da união de dois conjuntos, a qual é escrita como

$$n(A \cup B) = n(A) + n(B) - n(A \cap B). \tag{4.3}$$

Dividindo esta equação por $n(\Omega)$ e usando a abordagem clássica, identificamos a identidade

$$P(A \cup B) = P(A) + P(B) - P(A \cap B). \tag{4.4}$$

está equação é conhecida como regra da soma. O evento $A \cap B$ refere-se a ocorrência simultânea de A e B. Porém, se os eventos são mutualmente excludentes, $A \cap B = \emptyset$, então

$$P(A \cup B) = P(A) + P(B), \tag{4.5}$$

concordando com o terceiro axioma de Komolgorov.

Exemplo 4.2.2. Calcule, através da regra da soma, a probabilidade de ocorrer faces com números pares ou primos no lançamento de um dado "honesto".

Solução: Seja A o evento do lançamento do dado em que ocorre pares, $A = \{2, 4, 6\}$, e B o evento de ocorrência de números primos, $B = \{2, 3, 5\}$. Para calcular $P(A \cup B)$ através da regra da soma, precisamos achar $P(A)$, $P(B)$ e $P(A \cap B)$. Já sabemos que o número de elementos do espaço amostral no lançamento de um dado é $n(\Omega) = 6$. Portanto,

$$P(A) = \frac{n(A)}{n(\Omega)} = \frac{3}{6} = \frac{1}{2}$$

e

$$P(B) = \frac{n(B)}{n(\Omega)} = \frac{3}{6} = \frac{1}{2}.$$

Note que $A \cap B = \{2\}$ e, consequentemente,

$$P(A \cap B) = \frac{n(A \cap B)}{n(\Omega)} = \frac{1}{6}.$$

Finalmente, podemos usar a regra da soma

$$P(A \cup B) = P(A) + P(B) - P(A \cap B) = \frac{1}{2} + \frac{1}{2} - \frac{1}{6} = \frac{5}{6},$$

que é, justamente, a probabilidade de ocorrer A ou B.

4.2.2 Regra da multiplicação

Qual a probabilidade de acontecer os eventos A e B, considerando que estes são independentes? Para responder este questionamento, considere que Ω_A (Ω_B) é o espaço amostral que contém o evento A (B), ou seja,

$$A \subset \Omega_A, \qquad B \subset \Omega_B.$$

Estamos interessados no resultado geral proveniente do resultado individual de dois experimentos, cada um representado por seu espaço amostral Ω_A e Ω_B. Sendo assim, o espaço amostral geral pode ser escrito como o produto cartesiano* dos dois espaços amostrais individuais

$$\Omega \equiv \Omega_A \times \Omega_B = \{(a, b) \mid a \in \Omega_A \text{ e } b \in \Omega_B\}.$$

Analogamente, o evento global observado no experimento geral que está contido no espaço amostral Ω também pode ser representado como um produto cartesiano dos dois eventos individuais

$$AB = A \times B.$$

Onde AB e o evento de ocorrência de A e B. Com isso, temos

$$P(AB) = \frac{n(AB)}{n(\Omega)} = \frac{n(A \times B)}{n(\Omega_A \times \Omega_B)} = \frac{n(A)n(B)}{n(\Omega_A)n(\Omega_B)}.$$

Com isso, temos

$$P(AB) = P(A)P(B). \tag{4.6}$$

Esta equação é conhecida como regra da multiplicação para eventos independentes.

Porém, como determinar $P(AB)$ se os dois eventos forem dependentes? Suponha que a ocorrência de B depende da ocorrência de A. Os produtos cartesianos escritos anteriormente não podem ser mais usados, pois os eventos possuem dependência. A representação via teoria de conjuntos desta dependência entre os eventos é através de uma relação**. Neste caso, a dedução de $P(AB)$ seria extensa e, sendo assim, optamos por não entrar em tantos detalhes. No final desta dedução obteríamos

$$P(AB) = P(A)P(B|A), \tag{4.7}$$

*Perceba que o espaço amostral do lançamento de dois dados representado pela eq. (4.2) é o produto cartesiano dos espaços amostrais do lançamento de cada dado.

**Lembra o que é uma relação? Você a estudou em matemática no ensino médio. Trata-se de um subconjunto de um produto cartesiano. Este subconjunto obedece a uma condição. Vale a pena consultar um livro do ensino médio de matemática e relembrar a teoria dos conjuntos.

onde $B|A$ é a ocorrência do evento B considerando que o evento A aconteceu. Por causa desta dependência $P(B|A)$ é denominada de probabilidade condicional. Note que se A e B a eq. (4.6) é válida e, consequentemente $P(B|A) = P(B)$. Por isso, a eq. (4.7) é a regra da multiplicação em sua forma mais geral.

Exemplo 4.2.3. Considere uma caixa com três cartas enumeradas por 1, 2, e 3. Suponha que acontecem dois experimentos aleatórios, onde um cego retira da caixa uma única carta, a qual não retorna a caixa.
(a) Calcule, através da regra da multiplicação, a probabilidade da carta 2 ser retirada e em seguida a carta 3 seja selecionada?
(b) Calcule a mesma probabilidade do item anterior, usando a abordagem clássica diretamente.

Solução:

(a) A probabilidade da carta 2 ser retirada dentre as 3 cartas é

$$P(A_2) = \frac{1}{3}.$$

Considerando que o evento de retirada da carta 2 foi realizado, sobram apenas 2 cartas: 1 e 3. Portando, a probabilidade de que a carta 3 seja retirada considerando que a carta 2 já saiu é

$$P(A_3|A_2) = \frac{1}{2}.$$

Usando a regra do produto, obtemos a resposta

$$P(A_2A_3) = P(A_2)P(A_3|A_2) = \frac{1}{3}\frac{1}{2} = \frac{1}{6}.$$

(b) O espaço amostral da retirada de uma única carta é

$$\Omega_1 = \{1, 2, 3\}.$$

No entanto, o espaço amostral global pode ser representado através da seguinte relação:

$$\begin{aligned}\Omega &= \{(i,j)|\ i,j \in \Omega_1 \text{ e } i \neq j\} \\ &= \{(1,2);(1,3);(2,1);(2,3);(3,1);(3,2)\}\end{aligned} \quad (4.8)$$

O evento que estamos interessados é

$$A_2A_3 = \{(2,3)\}.$$

Sendo assim, podemos obter o resultado

$$P(A_2A_3) = \frac{n(A_2A_3)}{n(\Omega)} = \frac{1}{6},$$

concordando com o resultado do item (a).

4.3 Variáveis aleatórias e densidade de probabilidade

Qualquer evento aleatório pode ser rotulado por um valor numérico, o qual é chamado de variável aleatória. Estas podem ser discretas ou contínuas.

Considere um experimento em que ocorre N eventos mutualmente exclusivos, $A_1, \ldots, A_N$, e que x_j seja a variável aleatória discreta que representa o evento A_j. Sendo assim, podemos representar o espaço amostralcom a seguinte notação

$$\Omega = \{x_1, \ldots, x_N\} = \{x_j\}_{j=1}^{N}$$

e, também, usar a notação $P_j \equiv P(A_j)$. Sendo assim, pelo segundo e terceiro axiomas de Komolgorov, temos

$$\sum_{j=1}^{N} P_j = 1. \quad (4.9)$$

Isto quer dizer que a soma da probabilidade de todas as variáveis aleatórias discretas deve ser igual à unidade. Podemos também imaginar um espaço amostral contínuo, por exemplo, imagine um intervalo que define o espaço amostral $\Omega \equiv [a, b] \in \mathbb{R}$. Sendo assim, nossa variável aleatória x está dentro deste intervalo, ou seja, $a \leq x \leq b$. Podemos dizer que a probabilidade de acontecer um valor entre x e $x+dx$ é $\rho(x)dx$. Análogo ao caso discreto, ao invés de somar todas as probabilidade de eventos possíveis, podemos realizar uma integração, onde

$$\int_a^b \rho(x)dx = 1.$$

Esta função $\rho(x)$ é chamada de densidade de probabilidade. Usando a regra da soma de forma generalizada, podemos encontrar a probabilidade de achar x no intervalo $[c, d] \subset [a, b]$ da seguinte forma

$$P(c \leq x \leq d) = \int_c^d \rho(x)dx$$

A densidade de probabilidade para variáveis discretas pode também ser definido através da função δ de Dirac [3]

$$\rho(x) = \sum_{j=1}^{N} P_j \delta(x - x_j). \tag{4.10}$$

Note que se integrarmos esta equação, obtemos

$$\int_a^b \rho(x)dx = \sum_{j=1}^{N} P_j \int_a^b \delta(x - x_j) = \sum_{j=1}^{N} P_j = 1.$$

4.4 Momentos

Seja x uma variável aleatória pertencente à um intervalo Ω. O valor médio de uma função desta variável é definido como sendo

$$\langle f(x) \rangle \equiv \int_{\Omega} f(x)\rho(x)dx \tag{4.11}$$

O k-ésimo momento de x é definido da seguinte forma

$$\langle x^k \rangle \equiv \int_\Omega x^k \rho(x) dx.$$

Com isso vemos que o primeiro momento (com $k = 1$) é, justamente, a média $\langle x \rangle$. O desvio quadrático médio, também chamado de variância, é definido por

$$\text{var}(x) \equiv \langle (x - < x >)^2 \rangle. \tag{4.12}$$

A variância pode ser usada como uma estimativa de quanto os eventos aleatórios se afastam de sua média, ou seja, uma medida de dispersão. Usando a propriedade

$$\langle f(x) + g(x) \rangle = \langle f(x) \rangle + \langle g(x) \rangle, \tag{4.13}$$

podemos demonstrar que

$$\text{var}(x) = \langle x^2 \rangle - \langle x \rangle^2. \tag{4.14}$$

Note que $\langle x^2 \rangle$ é o segundo momento, também conhecido como média quadrática.

4.5 Conclusão

Podemos considerar que esta aula oferece um conteúdo básico, porém essencial para o tratamento estatístico que iremos precisar no estudo da mecânica estatística.

4.6 Resumo

Nesta aula vimos alguns conceitos básicos de estatística. Abordamos probabilidade através da teoria de conjuntos. Vimos que a

probabilidade de um evento A acontecer é

$$P(A) = \frac{n(A)}{n(\Omega)},$$

onde Ω é o espaço amostral do experimento. Mostramos duas regras importantes: a da soma

$$P(A \cup B) = P(A) + P(B) - P(A \cap B),$$

e a da multiplicação

$$P(AB) = P(B)P(A|B).$$

Vimos que um problema com variável aleatória discreta pode ser mapeado em um problema de variável aleatória contínua, usando a δ de Dirac da seguinte forma

$$\rho(x) = \sum_{j=1}^{N} P_j \delta(x - x_j).$$

A condição de normalização da probabilidade deve ser satisfeita

$$\int_{\Omega} \rho(x)dx = 1 = \sum_{j=1}^{N} P_j.$$

A média de uma função de x é obtida usando a seguinte equação

$$\langle f(x) \rangle = \int_{\Omega} f(x)\rho(x)dx.$$

4.7 Atividades

ATIV. 4.1. Um sistema consiste de quatro componentes independentes e funciona somente quando todas as componentes funcionam. Seja S_i o evento da componente i funcionar e $p_i \equiv P(S_i)$. Qual a probabilidade do sistema falhar?

Comentário: Usando a regra da multiplicação é possível encontrar a probabilidade do sistema funcionar. Note que o evento do sistema falhar é complementar ao evento deste funcionar.

ATIV. 4.2. Sendo $B_1, B_2, \ldots, B_N$ eventos mutualmente exclusivos e exaustivos, isto é, $\Omega_B = B_1 \cup B_2 \cup \ldots \cup B_N$, demonstre o teorema da probabilidade total,

$$P(A) = \sum_{j=1}^{N} P(A|B_j)P(B_j),$$

onde A é o evento de um experimento distinto do qual os eventos $B_1, B_2, \ldots, B_N$ podem ser observados.

Comentário: Inicialmente, pode ser mais fácil mostrar que $P(A) = P(A\Omega_B)$. Depois disso, que tal decompor Ω_B na união descrita no enunciado e, em seguida, usar as regras da multiplicação e da soma?

ATIV. 4.3. Um teste de diagnóstico de câncer tem 95% de precisão, tanto nos portadores do mal quanto nos não-portadores. Se 0,6% da população tem câncer, (a) calcule a probabilidade de determinado indivíduo ser portador do mal, sabendo-se que o resultado do teste foi positivo. (b) Se fizessem mais dois testes de diagnóstico e o resultado fosse positivo em ambos, qual seria a probabilidade do indivíduo ter câncer?

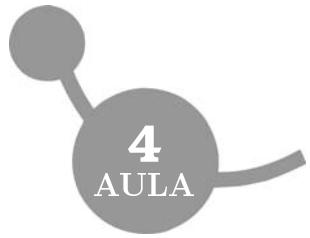
Comentário: Esta atividade requer alguns cuidados. É possível resolvê-la facilmente através do método bayesiano [2]. Uma consulta na ref. [5], poderá auxiliar sua resolução.

ATIV. 4.4. Em um programa de auditório, é realizado um jogo com três portas fechadas e um prêmio atrás de uma delas. Para ganhá-lo, o concorrente deve adivinhar a porta onde o prêmio está. O apresentador do programa convida uma pessoa do auditório, João, para concorrer ao prêmio, escolhendo uma das portas. (a) Qual a probabilidade de João ganhar o prêmio? Depois de algum suspense, o apresentador abre uma das portas que não contem o prêmio. Depois deste fato, ele oferece a João a opção de trocar de porta. (b) Do ponto de vista estatístico, vale a pena João realizar esta troca? Por que? Justifique sua resposta calculando probabilidades.*

*Este é um problema tradicional de probabilidade conhecido como problema de Monty Hall.

Comentário: Este problema pode ser resolvido de maneira análoga à atividade anterior. Uma leitura na ref. [4], facilitará sua resolução.

ATIV. 4.5. Demonstre as eqs. (4.13) e (4.14).

Comentário: A soma de funções é uma função também. Que tal usar a eq. (4.11)?

ATIV. 4.6. Partindo da eq. (4.11), mostre que

$$\langle f(x)\rangle = \sum_{j=1}^{N} f_j P_j,$$

onde $f_j \equiv f(x_j)$.

Comentário: É possível usar a eq. (4.10) e as propriedades da função δ de Dirac [3].

ATIV. 4.7. Considere que $j = 1, 2, 3, 4, 5$ ou 6 é uma variável aleatória referente às faces de um dado. (a) Calcule a probabilidade P_j da face j cair voltada para cima no lançamento de um dado "honesto". Usando o resultado da atividade anterior, faça $f(x) = x$ e, em seguida, $f(x) = x^2$ para obter (b) $\langle x\rangle$ e $\mathrm{var}(x)$.

Comentário: Reveja o exemplo 4.2.1 . Lembre-se que a variância pode ser calculada através da eq. (4.14).

ATIV. 4.8. Escreva a densidade de probabilidade, $\rho(x)$, para o lançamento de um dado "honesto", onde os valores possíveis para x são 1, 2, 3, 4, 5 e 6, referentes a cada face do dado. Calcule $\langle x\rangle$ e $\mathrm{var}(x)$ usando a eq. (4.11) e compare com o resultado da atividade anterior.

Comentário: Considerando que você já calculou P_j na atividade anterior, que tal usar a eq. (4.10)?

ATIV. 4.9. Considere $x \in [-\frac{a}{2}, \frac{a}{2}]$ uma variável aleatória contínua. Suponha que qualquer valor neste intervalo tenha mesma chance de ocorrência. (a) Obtenha a densidade de probabilidade* da variável x. (b) Calcule $\langle x \rangle$ e $\text{var}(x)$.

*Nesta atividade, $\rho(x)$ é chamada de distribuição uniforme.

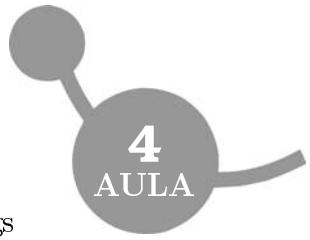
Comentário: Se os valores possuem a mesma chance de ocorrer, a densidade de probabilidade deve ser constante. Lembre-se da condição de normalização desta densidade.

ATIV. 4.10. A variável aleatória $x \in \mathbb{R}$ é regida pela densidade de probabilidade $\rho(x) = a \exp[-b(x-c)^2]$, onde a, b e c são parâmetros reais. Determine a, b e c em função de $\langle x \rangle$ e do desvio padrão de x, definido como $\sigma \equiv \sqrt{\text{var}(x)}$. Reescreva $\rho(x)$ apenas usando $\langle x \rangle$ e σ como parâmetros**.

**Nesta atividade, $\rho(x)$ é chamada de distribuição normal ou gaussiana.

Comentário: Lembre-se da condição de normalização que toda densidade de probabilidade deve obedecer.

4.8 Próxima aula

Depois de estudarmos termodinâmica e alguns conceitos de estatística, estamos preparados para introduzir a mecânica estatística. É isto que faremos na próxima aula, tratando sistemas clássicos.

Referências

[1] TRIOLA, M. F. *Elementary Statistics*. 3.ed. Benjamin/Cummings Publishing Company, 1986.

[2] JAYNES, E. T. *Probability Theory*: The Logic of Science. 1.ed. Cambridge University Press, 2003.

[3] *Delta Function*: from Wolfram MathWorld. Disponível em: < http: //mathworld.wolfram.com/DeltaFunction.html>. Acesso em: 15 de abril de 2011.

[4] *Monty Hall problem*: from Wikipedia, the free encyclopedia. Disponível em: < http://en.wikipedia.org/wiki/Monty_Hall_problem>. Acesso em: 18 de abril de 2011.

[5] PENA, S. D. Thomas Bayes: o cara! *Ciência Hoje*. v. 38, p. 22, jul. 2006.